Gaze and Task Performance in Shared Virtual Environments[1]

Jeremy N. Bailenson, Andrew C. Beall, & Jim Blascovich

Research Center for Virtual Environments and Behavior,

University of California, Santa Barbara

Jeremy N. Bailenson

Department of Psychology

University of California

Santa Barbara, CA 93106-9660

bailenson@psych.ucsb.edu

Abstract

Nonverbal behavior, particularly gaze direction, plays a crucial function in regulating conversations and providing critical social information. In the current set of studies, we represented interactants in a shared immersive virtual environment (IVE). Interactants sat in physically remote rooms, entered a common virtual room and played games of 20 questions. The interactants were represented by one of three types of avatars 1) human forms with head movements rendered in real-time, 2) human forms without head movements rendered, or 3) human voice only (i.e., a conference call). The data demonstrated that interactants in the rendered head movement condition rated a higher level of copresence, liked each other more, looked at each other's heads more, and spoke for a lower percentage of time during the game, compared to the other two conditions. We discuss implications for the design of shared virtual environments, the study of nonverbal behavior, and the goal of facilitating efficient task performance.

Introduction

Normal face-to-face communication is an extremely rich, multi-modal form of expression. Aside from the verbal channels, non-verbal channels available during face-to-face communication include gaze from head posture and eye direction, arm gestures, body posture, and facial expressions (as well as non-verbal aspects of language such as variations in intonation and voice quality). Surely if these other channels did not offer unique or additional "side-channel" information carrying capacity, then telephony-based conferencing would then be considered by most users to be equivalent to face-to-face meetings. Since this is not the case, researchers seek to discover the individual capacities of the different channels in order to better understand face-to-face meetings. One way to view the effect of multiple simultaneous channels is that of increasing the overall bandwidth available to the persons involved. For instance, adding a video component to phone conversation should theoretically increase the throughput of communication between participants. From an information standpoint this is certainly possible, but for the human participants it may not be possible to take advantage of the added information potential for a number of reasons. If, however, humans do use the additional channel, then two outcomes are possible: 1) more total information throughput may be achieved by using both channels, or 2) constant information throughput may be maintained by reducing the load on the original channel onto the new channel.

The use of immersive virtual environments (IVEs) to communicate and interact is becoming more common. Consequently, researchers are beginning to study the design and use of these environments. In particular, there are many studies that address the importance of rendering nonverbal behaviors in IVEs. Along the same lines, social scientists are beginning to use virtual environments as a tool to study human behavior, specifically nonverbal behaviors. In this current work, we are most interested in the non-verbal cues afforded by head movements in collaborative IVEs.

Head movements are an important source of information for several reasons. First, head movements, like eye movements, are highly correlated with an individual's focus of attention. In fact, we argue that most often the two are generally pointed in the same direction. An obvious exception to this is when one is "stealing a glance" and does not want the focus of attention to be easily detected. While head and eye directions can be easily de-coupled from one another, so can eye direction be de-coupled from the actual focus of visual attention. This occurs when one fixates a visual target and yet moves the "spotlight" of attention away from the fovea of the eyes. While this is possible, most often eye direction is highly indicative of a person’s focus of attention, just as head direction is usually indicative of eye direction. Given this tight coupling, we believe that head direction can be used by others to reliably monitor one's focus of attention. Head direction also offers distinct advantages as a source of this type of information when considering one's full field-of-view. Namely, it is possible to judge the head direction of another person even at retinal eccentricities of 90 deg, a location in the visual field which is well beyond the range which eye directions can be detected at all [1]. This is then a useful and valuable cue for detecting the attentional focus of an individual in one's far periphery. A second important source of information that head movements can contribute is symbolic information, as in indicating agreement, disagreement, and many other semantic messages [2,3,4]. For the above scientific reasons and for technical reasons concerning the challenge of accurately tracking eye movements in IVEs, this work uses avatars in which the head and eye directions are always locked together.

Gaining a better understanding of how gaze behavior communicates both a person's attention and intentions is an interest shared by researchers across a number of disciplines, including: computer scientists designing better IVEs, psychologists and studying gesture and communication, and organizational behavior researchers studying task performance. We review each of these bodies of literature in turn.

Gaze and Immersive Virtual Environments

Rendering gaze, head orientation, and other nonverbal gestures may be critical in order to produce a functional collaborative IVE [4,5,6,7]. However, doing so has its technological and pragmatic difficulties. Consequently, in this study we seek to measure the degree of improvement that rendering head orientation has on task performance. Furthermore, we can examine the degree of presence and copresence that conversationalists experience in an environment when rendered head movements are not available. Then, we can compare that amount of presence to a similar condition where the users have no visual representation (i.e., a conference call).

There is a large body of literature that addresses the issue of presence in virtual environments [8]. A subset of this literature deals with the concept of copresence. Copresence, also called social presence [9], occurs when a person uses a virtual environment that contains one or more virtual agents and behaves as if he or she were interacting with other veritable human beings. In other words, when a virtual agent influences a user of an IVE in a similar manner that a real person would, that user experiences high copresence.

A majority of the research on copresence inserts a user into an environment and asks him or her to answer questionnaires. While the information one gets from a questionnaire is clearly valuable, we believe that behavioral measures may offer the most reliable manner of measuring copresence, since they are less susceptible to demand characteristics (i.e., the person filling out the questionnaire simply reports what he or she thinks the experimenter wants to hear). For example, in previous work [10, 11], subjects walked around in a virtual room with an agent. As a behavioral measure of copresence, we measured the amount of personal space that our subjects left between themselves and the agent. We demonstrated that subjects left larger bubbles of personal space around agents that gazed realistically, compared to agents that did not gaze realistically. Consequently, proxemic behavior proved to be a reliable measure of copresence. In the current work, we seek to demonstrate (also with behavioral measures) the contribution of real-time rendered head movements to copresence.

<u>Conversation/Gaze literature</u>

Psychologists have studied gaze for decades. However, using IVEs, we can manipulate expressed gaze in ways that before were not possible[2]. We can have conversationalists in a virtual room, and with the press of a button we can 'lock' their gaze, such that the other people in the room can see each other, but can no longer see each other's head movements. Consequently, we can assess the unique contribution of knowing where your conversational partner is looking (i.e., the orientation of his or her head) on interaction and on collaborative task performance. In other words, using IVEs we can have someone be present in a room, have that person be able to look around the environment, but not have his conversational partners be able to follow his head direction (despite the fact that the partner can see his face and eyes).

Kendon [3] argues that gaze regulates talking and listening behavior, signaling turn-changes and other important information concerning the interaction. Furthermore, these patterns of behavior change with the structure and content of conversation, in that a conversationalist can attempt to regulate and change his partner's conversational behavior. Specifically, one needs to check the direction in which the other conversationalist is looking. Based on this information, one can see if interactants are talking too much, too little, or conveying inappropriate information. During interactions, one uses gaze to gauge a listener's attention as well as his or her level of comprehension. Kendon [3] describes 'interactional synchrony', where the listener tailors his movements and linguistic behaviors to fit with those of the speaker, including head movements and body postures. Furthermore, gaze in social contexts serves to provide information, regulate interaction, express intimacy, exercise social control, and facilitate task goals [12].

In a virtual environment, if head movements and looking directions are not rendered, then we would expect a lower degree of interactional synchrony, since one cannot use those nonverbal cues to tailor his or her behaviors to the head movements of others. Consequently, we would expect task performance to be more difficult in situations where those cues are absent, compared to situations where full gaze behavior is rendered, even if the task is an ostensibly purely verbal task. In addition, this lack of synchrony should decrease copresence as well.

Organizational/Task Performance Literature

Gaze also serves to facilitate the learning process and enhance task performance. During instruction, gaze helps learning, in that college students had higher performance on a learning task when the instructor gazed at them than when the instructor did not [13]. Furthermore, when students are able to return the gaze to the instructor, they participate more in the instruction than when they are not able to gaze at the instructor [14]. Situations that foster mutual gaze are especially preferred by interactants during cooperation tasks when multiple individuals are working towards a goal [15]. Finally, Short, Williams, & Christie [16] suggest that in order to see an effect of gaze on problem solving performance, the task needs to be interactive.

Overview of Experiment

In the current study, we seek to explore an additional behavioral measure of copresence. We examine the importance of knowing the exact direction that an avatar’s head is facing. In this experiment, our participants played the game of "twenty questions." This is a question-and-answer game in which one person (the answerer) answers questions from others (the interrogators) with simple "yes" or "no" responses. We use the participants' performance in the game as a measure of the importance of gaze. Furthermore, we measure their head movements, and use their looking behavior as an additional indicator of copresence.

Hypotheses

1. Avatars with rendered head movements should elicit more copresence than avatars without rendered head movements.

2. There should be more use of horizontal head movements (and less use of vertical head movements) when head movements are rendered, since looking at other players gaze behavior is most useful in this condition.
3. Task performance should be most effective when interactants are made aware of each others' head movements and interactional synchrony is high.

Methods

Design

The primary independent variable of interest was avatar behavior. There were three levels to this factor. In the low behavior condition, there was no visual representation for any avatar. Participants were immersed in a virtual room, but communicated verbally as they would during a speakerphone conference call. In the medium behavior condition, each participant could see the other two participants' avatars and hear their voices. Furthermore, the avatars blinked, and when a given participant spoke, his or her avatar's mouth opened and closed in close synchrony to the amplitude profile of the speech intensity (we refer to this cue as lip-flapping). In the high behavior condition, the avatar blinked and lip-flapped. However, in this condition, we also rendered participants' head movements. Consequently, at any given time, it was possible for one participant to see the directions in which the other two participants were looking.

This behavioral variable was varied within subject. Each participant played three separate blocks of 20 questions, one for all three levels of behavior. They played three games in each block, resulting in nine games total. Order of blocks was counterbalanced such that each behavior type appeared in each block position an equal number of times. For each of the nine games of 20 questions, participants attempted to identify a different word. Order of words was counterbalanced such that each word appeared in each serial position an equal number of times. Furthermore, each word appeared in each behavior condition the same number of times.

Materials and Apparatus

The equipment used to render our virtual environments is described in detail in Bailenson, Blascovich, Beall and Loomis [10]. The head mounted displays (HMDs) were Virtual Research V8 HMDs with 680 by 480 resolution stereoscopic LCD displays running at 72 Hz refresh rates. Visual rendering was stereoscopic with a 60 degrees diagonal field of view. The display updated on average at a frame rate of 36 Hz with a latency of less than 65 ms between subjects' head movements and the concomitant update in the visual display. Participants' orientation was tracked using inertial tracking. The computer used to run the experiment was a 450 MHz dual-processor, Intel Pentium III. The graphics were produced using Evans & Sutherland Tornado 3000 video-cards.

Participants wore microphones and auditory headsets over their HMDs. Consequently, they could hear each other speak. Furthermore, the experimenter could speak into a microphone and address them individually (e.g., to tell the answerer the word for the next trial) or as a group (e.g., to frequently check for any symptoms of simulator sickness). We used custom, real-time audio sampling software to measure the instantaneous speech sound levels captured near each participant's mouth. When the amplitude was over a certain threshold, their avatars opened their mouths to indicate speech. We sampled the microphone amplitude at a rate of 20 Hz.

Participants

Participants were 27 undergraduate students (16 male, 11 female) who received partial credit in an introductory psychology course for participation.

Procedure

There were three participants in each session. Each participant sat in a separate room, where they were not in physical sight of each other. In each room, they put on an HMD. At that point, they entered the same virtual room. Panel A of Figure 1 shows the arrangement used in the experiment. Each IVE facility was networked together to bring the three participants together in the virtually shared room around a common table. Panel B of the figure shows an outsider's view (i.e., not the viewpoint of one of the three participants) of the virtual room and avatars. We used two different female faces and two different male faces, and attempted to match them to resemble our participants. Participants could not see their own avatar.

We instructed participants that they would be playing a game of 20 questions. We chose this task because it can be construed as a purely verbal task, and has been utilized in previous reasoning studies [17]. We informed participants that sometimes they would be able to see representations of the other players while at other times they would only be able to hear them. One of the three participants was always the anwerer, while the other two interrogators asked ‘yes or no’ questions to determine the answer. We used three players in each game (as opposed to two players) in order to maximize the amount of head movements among players. As panel B of Figure 1 demonstrates, it was necessary for players to move their heads in order to look from one player to another.

In each condition, each participant was the answerer in one of the three trials and was the interrogator in the other two trials. The answer was always some kind of an object (i.e., Boat, Airplane, Rock, Ocean, Frog, Monkey, Carrot, Apple, Chair). We instructed the interrogators to figure out the answer in as few turns as possible, and that they did not necessarily need to alternate turns when asking questions.

After each of the three blocks, participants took off their HMDs to rest and to fill out the questionnaire shown in Appendix I. They indicated their agreement on a seven point Likert scale, with higher numbers indicating more agreement. After the nine games were complete, we collected biographical information about participants, including their age, gender, religiosity, virtual reality experience, and frequency of video game usage.

Results

We looked at a number of dependent variables. First we will discuss behavioral results from performance in the game itself. Next we examine the ratings data from the questionnaires.

Table 1 shows the means and standard deviations of the number of questions that participants asked per game, the average time it took for a participant to ask a question, average time it took for participants to finish each game by condition, and the average time individual participants spent speaking during the game. Two experimental assistants coded subjects' behavior as they asked questions to determine whether or not an utterance qualified as a question. Interobserver reliability was high (r = .95) and differences were resolved by listening to recordings after the sessions had ended.

As the table shows, there was a (non-significant) trend for participants to need fewest questions to finish the game in the high behavior condition (M = 8.73) than in the other two conditions (M = 9.80) The next measure we examine is the proportion of time that participants spoke in each of our behavior conditions. From the recorded voice amplitudes levels and the criterion threshold, we calculated the percentage of time that each participant spoke during the game. We found a monotonic trend, with the highest percentage of speaking in the low behavior condition (M = 7.56%), less in the medium behavior condition (M = 6.60%,), and the least in the high behavior condition (M = 6.15%), $F(1,26)=8.14$, $p<.01$. However, while the proportion of speaking decreased, the total amount of speaking (measured in time) was the same for the three conditions, indicating that there were more periods of silence in the high behavior condition. Table 1 also indicates a nonsignificant trend for games in the high behavior condition to take longer than games in the other conditions. However, it is difficult to draw inferences from this trend since the instructed goal for our subjects was to finish using the fewest amount of questions, not the least amount of time.

The next behavioral measure we examine is head orientation, which we sampled at a rate of 15 Hz. We were most interested in movements in the horizontal plane, because these head rotations likely indicate instances in which participants looked towards the other players. We computed the average standard deviation of rotations for each condition (measured in degrees). Higher standard deviations indicate less time spent looking straight ahead. The most movement occurred in the high behavior condition (M = 8.91), followed by the low behavior condition (M = 7.82) and the medium behavior condition (M = 7.56). We were most interested in the comparison between the medium and high conditions, since higher standard deviations in the high condition would presumably indicate that participants were monitoring each other’s head movements. A paired t-test comparing these two conditions demonstrates near significance, $t(26)=2.03$, $p<.053$.

We ran a similar analysis examining movements in the vertical plane. There was higher standard deviation in the low behavior condition (M = 6.53) than in the medium (M = 3.62) or high (M = 4.06) behavior conditions, $F(2,25)=8.69$, $p<.001$. Post-hoc tests demonstrated no significant difference between the high and medium conditions. This effect makes sense; when there are visible avatars in the room, participants tend not to move their heads far from eye level thus reducing the amplitude of vertical motion. However, without the presence of avatars, people may be inclined to spend more time looking up and down.

For the questionnaire ratings, we computed the average score across the 30 questions (ten per condition). Cronbach's alpha, a common measure used in psychology to determine the reliability of a scale, across the ten questions was .84. A high score indicates the participant felt involved in the game, felt as if he or she was truly interacting with the other players, and enjoyed the experience. A low score indicates the opposite.

The average score across conditions was 3.86. We ran a within-subject ANOVA with condition as the independent variable and score as the dependent variable. There was a significant effect, $F(2,25)=3.93$, $p<.05$ with the highest score in the high behavior condition (M = 4.08, SD = .87), followed by the medium behavior condition (M = 3.68, SD = .87), and the low behavior condition (M = 3.77, SD = .75). This effect was driven mostly by question 5 ("I felt like I was really interacting with the other players in the game) and question 8 ("I like the other players in the game"). Independent ANOVAs on each of these questions by condition demonstrated significant effects, with highest scores in the high behavior condition. There were no significant differences on the other individual questions.

Discussion

In this work, the task was designed to require multiple people to work together and take advantage of the non-verbal communication channel of head orientation. The game of 20 questions was also a useful paradigm to promote a relatively free-form interaction between the three participants and lent itself to multiple replications within-subjects over a relatively short period of time. This work revealed several interesting findings in both the objective behavioral data and the subjective questionnaire data.

The behavioral data showed that when we enabled the transmission of head gaze information, participants responded in two ways. First, the overall pattern of head movements changed: the amount of motion in the horizontal plane increased while there was a trend for the amount of motion in the vertical plane to decrease. We attribute this pattern to two possibilities: either participants were glancing back and forth to monitor the attentional focus of the others, or the participants were signaling their attentional focus by pointing with their heads at either of the others. The current study cannot distinguish between these two possibilities.

The second way in which participants responded to rendered head movements was by reducing the overall proportion of time spent verbally communicating. While there was no overall difference between the total amount of talking between experimental conditions, when interactants were made aware of the head movements of others, the proportion of time spent talking decreased. This finding was contrary to our predictions and signifies that in our study, adding an extra channel of non-verbal communication may have actually diluted the amount of time dedicated to solving the problem.

The subjective data supports the conclusion that communication was enhanced. This is indicated by the significant shift in the positive direction of the questionnaire ratings, and more importantly by the fact that as the behavioral manipulation was changed from low to high, this positive shift was largely driven by questions #5 and #8. These two questions (and possibly question #4) were arguably the only questions out of the 10 that asked the participant about the interpersonal aspects of the experience. The other questions probed the more environmental and task related aspects of the participants' experience. Therefore, as indicated by the subjective responses, adding non-verbal channel led participants to both feel as though they were interacting more with others and that they liked the other participants more. However, when coupled with the speaking percentage data, it seems that their enjoyment and copresence may come at the expense of focusing on the task at hand, since rendering the head gaze did not significantly improve task performance but may have diluted the interaction.

Conclusions. In summary, these findings confirmed our expectation that participants engaged in a collaborative task would readily take advantage of the additional communication channel afforded by head orientation. As discussed in the introduction, there are numerous findings in several fields that readily point to the prediction that even a seemingly verbally oriented task—the game of 20 questions—would benefit from the inclusion of simple non-verbal cues. We do not demonstrate this type of benefit from rendering head movements. However, this may be do to a floor effect, in that the task was too simple to demonstrate any improvements that may result from the knowledge of other interactants' head orientation. Indeed, in this study we find some evidence that task performance could possibly suffer as a result of including an extra nonverbal channel.

In future studies, we plan to expand on this paradigm by examining more complicated tasks as well as and testing other non-verbal channels of communication beyond head pose, namely eye gaze, arm gestures, postures, and facial expressions. Given the current data, designers of shared virtual environments [19, 20] can accrue some of the subjective enjoyment of an interaction by merely rendering head movements. While tracking eye movements surely will add to the copresence and the amount of conversational cues governing the interaction, the cost and difficulty of doing so may not currently be advantageous.

Table 1. Average number of questions asked by a single participant per game, time taken per question, and total time per game in each condition. Time is measured in seconds, and corresponding standard deviations are in parenthesis.

	Avatar Behavior		
	Low	Medium	High
Number of questions	9.82	9.77	8.73
	(4.23)	(5.87)	(4.56)
Time per question	7.20	8.01	7.99
	(2.80)	(3.30)	(1.64)
Time per game	144.11	162.52	192.97
	(74.21)	(91.69)	(118.11)
Speaking Percentage	7.56	6.60	6.15
	(0.05)	(0.05)	(0.04)

Footnotes

1. The authors would like to thank Eyal Aharoni and Jack Loomis for helpful comments on this paper as well as Meg Brzezinska and Mike Raimundo for assistance in collecting data. This research was sponsored in part by NSF Award SBE-9873432.

2. Although, see Argyle, Lalljee & Cook [18] for an example of segmenting components of mutual gaze without the aid of digital environments.

Appendix I. Items from the questionnaire

_____ 1. It was a lot of fun

_____ 2. I felt like I was really in a room.

_____ 3. Playing in virtual reality was harder than it would be to play the same game in a physical room.

_____ 4. I would like to spend more time with the other players in virtual reality.

_____ 5. I really felt like I was interacting with the other people.

_____ 6. After a while, it felt like the virtual room was a real room.

_____ 7. I was motivated to be involved with the game.

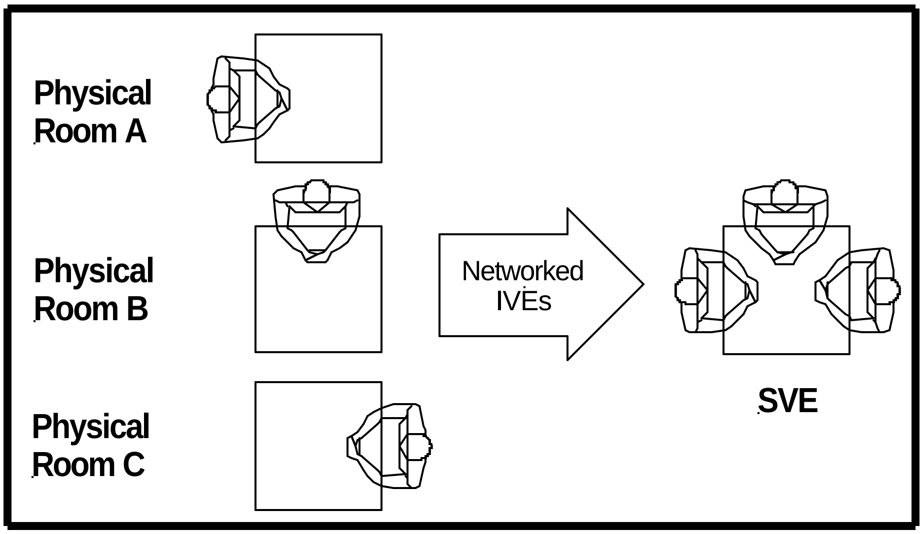
_____ 8. I liked the other players of the game.

_____ 9. Playing 20 questions is easy.

_____ 10. The other two players are good at playing 20 questions.

Figure 1. The configuration of the virtual room.

Panel A



Panel B

References

1. Pusch, M. & Loomis, J. M. (2001). Judging another person's facing direction using peripheral vision. Journal of Vision, 1.
2. Argyle, M. & Dean, J. (1965). Eye contact, distance, and affiliation. Sociometry, 28,l 289-304.
3. Kendon, A. (1977). Studies in the Behavior of Social Interaction. Indiana University: Bloomington.
4. Vertegaal, R., Slagter, R., van der Veer, G., & Nijholt, A. (2000). Eye gaze patterns in conversations: There is more to conversational agents than meets the eyes. SIGCHI, '01, Seattle, WA.
5. Cassel, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H., and Yan, H. (1999). Embodiment in conversational interfaces: Rea. In Proceedings of CHI'99, pages 520-527. Association for Computing Machinery, ACM Press.
6. Perlin, K. & Goldberg, A. (1996). Improv: A System for Scripting Interactive Actors in Virtual Worlds. Computer Graphics; Vol. 29 No. 3.
7. Normand, V., Babski, C., Benford, S., Bullock, A., Carion, S., Farcet, N., Frecon, E., J.Harvey, N. Kuijpers, N. Magnenat-Thalmann, S. Raupp-Musse,T.Rodden, M. Slater, G. Smith, A. Steed, D. Thalmann, J. Tromp, M. Usoh,G.Van Liempd, N. Kladias, 1999. TheCOVENproject: exploring applicative, technical and usage dimensions of collaborativevirtual environments, Presence: teleoperators and virtual environments, Vol.8, No2, pp.218-236.

8. Durlach, N. & Slater, M. (2000). Presence in shared virtual environments and virtual togetherness. PRESENCE, 9, 214-217.

9. Blascovich, J., Loomis, J.M., Beall, A.C., Swinth, K., Hoyt, C., & Bailenson, J. (in press). Using Immersive Virtual Environment Technology as a Methodological Tool in Social Psychology. Psychological Inquiry.

10. Bailenson, J.N., Beall, A.C., Blascovich, J., Weisbuch, M., & Raimmundo, R. (2001). Intelligent Agents Who Wear Your Face: Users' Reactions to the Virtual Self. Lecture Notes in Artificial Intelligence, 2190, 86-99.

11. Bailenson, J.N., Blascovich, J., Beall, A.C., & Loomis, J. (2001). Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. Presence, 10 583-597.

12. Kleinke, C. L. (1986). Gaze and eye contact: A research review. Psychological Bulletin, 100, 78-100.

13. Fry, R. & Smith, G. F. (1975). The effects of feedback and eye contact on performance of a digit-encoding task. Journal of Social Psychology, 96, 145-146.

14. Caproni, V., Levine, D., O'Neal, E., McDonald, R., & Garwood, G. (1977). Seating position, instructor's eye contact availability, and student participation in a small seminar. Journal of Social Psychology, 103, 315-316.

15. Jellison, J.M., & Ickes, W.J. (1974). The power of the glance: Desire to see and be seen in cooperative and competitive situations. Journal of Experimental Social Psychology, 10, 444-450.

16. Short, J., Williams, E., & Christie, B. (1976). The Social Psychology of Telecommunications. Chichester: Wiley.

17. Kafer, K. L.& Hunter, M. (1997). On testing the face validity of planning/problem-solving tasks in a normal population. Journal of the International Neuropsychological Society. 3, 108-119.

18. Argyle, M., Lalljee, M. & Cook, M. (1968). The effects of visibility on interaction in a dyad. Human Relations, 21, 3-17.

19. Coburn, R. A., Cohen M. F., & Drucker, S. M. (2000). The role of eye gaze in avatar mediated conversational interfaces. Microsoft Research Technical Report.

20. Johnson, W. L., Rickel, J. W., Lester, J. C. (2000). Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. International Journal of Artificial Intelligence in Education 11:47-78.